

(19) 世界知的所有権機関
国際事務局



(43) 国際公開日
2002 年1 月31 日 (31.01.2002)

PCT

(10) 国際公開番号
WO 02/08911 A1

(51) 国際特許分類: G06F 12/12, 12/08, 12/10,
9/06, 9/45, G10K 5/02, H04N 5/907

(SAITOU, Yasuhiko) [JP/JP]. 海永正博 (KAINAGA, Masahiro) [JP/JP]; 〒187-8588 東京都小平市上水本町五丁目20番1号 株式会社 日立製作所 半導体グループ内 Tokyo (JP).

(21) 国際出願番号: PCT/JP00/04906

(22) 国際出願日: 2000 年7 月24 日 (24.07.2000)

(74) 代理人: 玉村静世 (TAMAMURA, Shizuyo); 〒271-0092 千葉県松戸市松戸1333番地1 コスモ松戸ステーションスクエア209号 Chiba (JP).

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(81) 指定国 (国内): JP, KR, US.

(71) 出願人 (米国を除く全ての指定国について): 株式会社 日立製作所 (HITACHI, LTD) [JP/JP]; 〒101-8010 東京都千代田区神田駿河台四丁目6番地 Tokyo (JP).

添付公開書類:
— 国際調査報告書

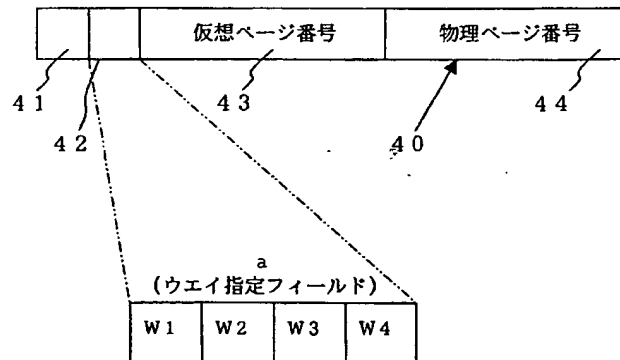
(72) 発明者; および

2 文字コード及び他の略語については、定期発行される各 PCT ガゼットの巻頭に掲載されている「コードと略語のガイダンスノート」を参照。

(75) 発明者/出願人 (米国についてのみ): 斎藤靖彦

(54) Title: DATA PROCESSING SYSTEM

(54) 発明の名称: データ処理システム



43...VIRTUAL PAGE NUMBER

44...PHYSICAL PAGE NUMBER

a...(WAY SPECIFYING FIELD)

(57) Abstract: A data processing system comprises a page information table (15) for holding page information on a page constituting an address space of a program to be executed and a cache memory (5) of multi-way type. A way specifying field (42) for specifying a way of the cache memory is provided in each page information table (40) in the page information table. The cache memory selects a cache fill object way from the ways specified by the way specifying fields corresponding to the page of the access address in replacing stored information in response to a cache miss. Thus, a way specifying field can be so set that the way of storage destination of the page containing information of high spatial reusability may be different from that of the page containing information of low reusability. As a result, in replacing stored information in response to a cache miss, a failure that information of relatively high reusability is expelled from the cache memory by information of relatively low reusability.

[続葉有]



(57) 要約:

データ処理システムは、実行すべきプログラムのアドレス空間を構成するページのページ情報を保持するページ情報テーブル（１５）と、マルチウェイ形式のキャッシュメモリ（５）とを含む。ページ情報テーブル内の各ページ情報フィールド（４０）には、キャッシュメモリのウェイを指定するウェイ指定フィールド（４２）を設ける。キャッシュメモリは、キャッシュミスに応答する記憶情報のリブレースにおいて、アクセスアドレスのページに対応するウェイ指定フィールドによって指定されるウェイの中からキャッシュフィル対象ウェイを選択する。これにより、空間的に再利用性の高い情報を持つページと再利用性の低い情報を持つページとの格納先ウェイが異なるようにウェイ指定フィールドを設定することにより、キャッシュミスに起因する記憶情報のリブレースにおいて、再利用性の比較的高い情報が再利用性の比較的低い情報によってキャッシュメモリから追出される事態を低減できる。

明 細 書

データ処理システム

5 技術分野

本発明は仮想記憶及びマルチウェイ形式のキャッシュメモリを有するデータ処理システムやデータプロセッサに関し、更には、そのようなデータ処理システムやデータプロセッサの為のプログラム開発方法、そのようなデータ処理システムやデータプロセッサの為のプログラムを記録した記録媒体に関する。

背景技術

データプロセッサの動作周波数と、データプロセッサとメインメモリを結ぶ外部バスの動作周波数との格差は年々増大する傾向にある。こうしたシステム構成ではメインメモリから処理に必要なデータやプログラムをデータプロセッサ内に持ってくるまで、プロセッサの側の処理が止まり、処理時間が増大する場合がある。

これを解決するためにメインメモリよりも高速なメモリをデータプロセッサ内部にもち、頻繁に使用するデータやプログラムをこのデータプロセッサ内部のメモリに貯えて、データプロセッサの動作クロックと同等または数クロック程度の遅れでデータ及び命令を演算装置に供給可能にする。そのようなメモリがキャッシュメモリであり、さまざまなデータプロセッサで採用されている。

しかし、このキャッシュメモリは高速である反面、チップ面積の制約のためメモリ容量をメインメモリ程大きくできない。したがって、処理に必要なすべてのデータおよび命令をキャッシュメモリに保持させる

ことは一般的にできない。このためプログラムの実行にともない動的にキャッシュメモリ内の内容をメインメモリに格納されているデータおよび命令と置き換える必要がある。この置き換え（キャッシュエントリのリプレース）方式や連想記憶形式などに応じて、キャッシュメモリは、

5 例えば、完全連想形式、マルチウェイ(Multi Way)形式、ダイレクトマップ(Direct Map)形式などがある。これらのキャッシュメモリについて記載された文献には、「コンピュータ オーガニゼーション アンド ハードウェア ソフトウェア インターフェース」の第7章(1994年、Morgan Kaufman 出版、D.A.Patterson, J.L.Hennessy 共著「COMPUTER

10 ORGANIZATION & DESIGN THE HARDWARE/SOFTWARE INTERFACE」の第7章「Large and Fast:Exploiting Memory Hierachy」)がある。

キャッシュメモリをもつデータプロセッサであっても、キャッシュメモリに存在しないデータおよび命令は、やはりメインメモリからもってくる必要がある。これはキャッシュミスによるメインメモリアクセスと

15 呼ばれ、キャッシュミスが頻発するとデータ処理性能を全体的に低下させることになる。たとえば、プログラム全体を通じて頻繁にアクセスされる命令列やデータなど(このような命令列やデータは時間的再利用性が高いとも呼ばれる)がキャッシュメモリに留まっていれば、メインメモリに対するアクセス頻度が少なくなるので性能劣化は少ないが、比較

20 的小さなキャッシュメモリでは、大規模なプログラムに係るデータや命令列にアクセスした場合、再利用性の高い命令列やデータがキャッシュからメインメモリに追い出される可能性が高くなり、性能低下の一因となる。このような観点より、キャッシュミスの低減が要請される。

特開平3-52044号公報には、キャッシュメモリ管理方式として、

25 仮想メモリ管理ユニットのページテーブルにページ単位でキャッシュ識別子を設け、与えられた仮想アドレスに対応するキャッシュ識別子に

従って複数キャッシュメモリの中から一つのキャッシュメモリを選択して利用する制御形式を開示する。要するにこの技術は、異なるキャッシュ置き換えアルゴリズムをもつ複数のキャッシュメモリを用意し、仮想アドレス空間を構成する各ページ対応で、キャッシュメモリを
5 対応させるものである。

画像や音声、信号処理などのマルチメディア処理で扱われるデータ量はキャッシュメモリ容量より大きい場合が多い。このようなマルチメディア処理分野のプログラムのデータアクセスの特徴として、第1に、個別のデータ要素は一度アクセスされると以後の処理で再び利用される機会が比較的少なく、第2に、アドレス上隣り合う各データ要素が時間的に連続してアクセスされる傾向が強い、ということがある。
10

これをキャッシュ上での各データ要素のアクセスの観点からみれば、マルチメディア処理では、キャッシュメモリ上のデータ列はアドレスが近ければ時間的に連続して利用できる、すなわち空間的再利用性は高いが、それらデータ列へのアクセスが一旦終われば、その後、キャッシュメモリにずっと留めておいても、同じデータを再び参照する機会が低い場合がある。すなわち時間的再利用性が低い。このようなデータ及びアクセスを夫々、ストリームデータ及びストリームアクセスと呼ぶ。
15

組み込み用途のプロセッサなどでは小さなチップ面積が要求されるため、キャッシュメモリを大きくすることに制約がある。そうすると、ストリームアクセスにより、キャッシュメモリにあるデータの置き換えが多発する傾向になる。そうなると、プログラム全体を通じて頻繁にアクセスされる時間的再利用性の高いデータも一緒にキャッシュメモリからメインメモリに追出される結果になる。従来技術ではこの点について考慮はなく、キャッシュミスを比較的多く引き起こす結果になってしまう。
20
25

また、頻繁にアクセスされる複数の命令列に対してもキャッシュメモリの記憶容量が小さければ一方の命令列の実行に片寄ったとき頻繁にアクセスされるもう一方の複数の命令列はメインメモリに追い出され、次に当該もう一方の命令列を頻繁に実行しようとする段階ではキャッシュミス

5 シュミスを頻発してしまう。

そこで、本発明者は、キャッシュミスに応答するキャッシュエントリのリプレースによって記憶情報の追い出しが起こる点に着目し、キャッシュエントリのリプレースに際してその対象を選択的に指定可能にすることの有用性について見出した。

10 本発明の目的は、キャッシュエントリのリプレースに際してその対象を選択的に指定可能なキャッシュ制御技術を提供することにある。

本発明の別の目的は、キャッシュエントリのリプレースに際して、時間的にも空間的にも再利用性の比較的高い情報が再利用性の比較的低い情報によって追い出される事態を抑制することが可能なキャッシュ

15 制御技術を提供することにある。

本発明の更に別の目的は、キャッシュエントリのリプレースに起因して生ずるキャッシュミスの増大を比較的に、しかも柔軟に、抑制することが可能なデータ処理システム及びデータプロセッサを提供することにある。

20 本発明の上記並びにその他の目的と新規な特徴は本明細書の以下の記述と添付図面から明らかにされるであろう。

発明の開示

〔1〕データ処理システムは、中央処理装置と、前記中央処理装置が実行するプログラムのアドレス空間を構成するページのページ情報を保持するページ情報テーブルと、マルチウェイ形式のキャッシュメモリと

25

を含む。前記ページ情報テーブル内の各ページ情報フィールドには、キャッシュメモリのウェイを指定するウェイ指定フィールドを設ける。前記キャッシュメモリは、キャッシュミスに応答する記憶情報のリプレースにおいて、アクセスアドレスのページに対応するウェイ指定フィールドによって指定されるウェイの中からキャッシュフィル対象ウェイを選択する。

上記より、例えば空間的に再利用性の高い情報を持つページと再利用性の低い情報を持つページとの格納先ウェイが異なるようにウェイ指定フィールドが設定されることにより、キャッシュミスに起因する記憶情報のリプレースにおいて、再利用性の比較的高い情報が再利用性の比較的低い情報によってキャッシュメモリから追出される事態を低減でき、これによってキャッシュミスの低減に寄与することができる。

前記ウェイ指定フィールドは、最大で全てのウェイを指定可能な複数ビットを有してよい。これにより、リプレースの対象となり得る候補の数を任意に選ぶことができ、選ばれた候補の中から1個のウェイをリプレースすることができる。したがって、プログラムによるアドレス空間の利用形態若しくは利用頻度などに即したキャッシュメモリの制御若しくは運用が可能になり、更にキャッシュミスの低減が可能になる。

望ましい形態としては、指定された複数個のウェイがリプレースの対象になるとき、その中から一つのウェイを決定する選択論理を持つとよい。選択論理はランダム論理であってもよいが、キャッシュミス低減の観点よりすれば、最も最近参照されていないウェイを選択するLRU選択論理を採用してもよい。

前記ページ情報テーブルは、前記中央処理装置によってアクセス可能なメインメモリに形成される。前記ページ情報テーブルに格納されているページ情報の一部をコピーとして保有する連想記憶形式のアドレス変換

バッファを更に有してよい。アクセスの高速化の観点よりすれば、前記キャッシュメモリは、前記中央処理装置及びアドレス変換バッファと同一の半導体基板に形成されるのがよい。セカンドキャッシュメモリなどを想定すれば、前記キャッシュメモリは、前記中央処理装置及びアドレス変換バッファと異なる半導体基板に形成されてもよい。

〔2〕中央処理装置と、前記中央処理装置のアドレス空間を管理する為のページ情報エントリの記憶領域を有するアドレス変換バッファと、マルチウェイ形式のキャッシュメモリとを含むデータプロセッサの観点では、前記ページ情報エントリの記憶領域には前記キャッシュメモリのキャッシュミスに
10 応答するキャッシュフィル対象ウェイの候補を指定するウェイ指定フィールドを有する。前記キャッシュメモリは、アクセスアドレスに対応するページ情報エントリに含まれる前記ウェイ指定フィールドの値を入力し、キャッシュミスに
15 応答する記憶情報のリブレースにおいて、入力した前記ウェイ指定フィールドが指定する候補の中からキャッシュフィル対象ウェイを決定する制御手段を有する。

このデータプロセッサにおいても上記同様に、例えば空間的に再利用性の高い情報を持つページと再利用性の低い情報を持つページとの格納先ウェイが異なるようにウェイ指定フィールドが設定されることにより、キャッシュミスに起因する記憶情報のリブレースにおいて、再利用性の比較的高い情報が再利用性の比較的低い情報によってキャッシュメモリから追出される事態を低減できる。

前記ウェイ指定フィールドは、キャッシュフィル対象ウェイの候補を複数指定可能な複数ビットを有してよい。これにより、リブレースの対象となり得る候補の数を任意に選ぶことができ、選ばれた候補の中から1個のウェイをリブレースすることができ、プログラムによるアドレス空間の利用形態若しくは利用頻度などに即したキャッシュメモリの制御若しくは運用が
25

可能になり、更にキャッシュミスの低減に寄与する。

望ましい形態としては、指定された複数個のウェイがリプレースの対象になるとき、その中から一つのウェイを決定する選択論理として、LRU選択論理を採用してもよい。

- 5 前記キャッシュメモリはデータ用と命令用に分離されてよい。データと命令では時間的並びに空間的な再利用性が相違する場合が多いから、夫々に即したウェイ指定による効果を達成することが容易になる。前記キャッシュメモリはデータと命令に共用されるものであってもよい。

- 10 〔3〕マルチウェイ形式のキャッシュメモリと中央処理装置を備えたデータプロセッサにより実行可能な形式のプログラムの生成を支援するプログラム開発方法の観点では、プログラムを構成する所定の情報を仮想アドレス空間上の所定の領域にまとめて配置することを指示するプログラミング言語の指示文により、プログラム実行時に前記所定の情報をキャッシュメモリのどのウェイに割り付けるかを指示するようにする。
- 15 これによりユーザは、データプロセッサのプログラム開発において、データやプログラムの利用形態若しくは性質に即応させてそれら情報に対するウェイへの割り当てを任意に決定できる、というキャッシュ制御の自由度を簡単に得ることができる。

- 20 前述の方法により、プログラムのアドレス空間を管理する仮想記憶とマルチウェイ形式のキャッシュメモリとを用いてデータ処理を行うためのプログラムが開発され、開発されたプログラムは、コンピュータ読取り可能に記録媒体に記録して提供することが可能である。このとき、プログラムは、アドレス空間上における前記プログラムの配置を規定する管理情報を有し、この管理情報の中に、プログラムを構成する所定の
- 25 情報をキャッシュメモリのどのウェイに割り付けるかを指示するウェイ指定情報が含まれる。このプログラムを用いることにより、前述の記

憶情報のリブレースに際して、時間的にも空間的にも再利用性の比較的高い情報が再利用性の比較的低い情報によって追い出され事態の防止を容易に実現できる。

5 図面の簡単な説明

第1図は本発明に係るデータ処理システムの概略を示すブロック図である。

第2図はデータプロセッサの詳細を例示するブロック図である。

第3図はキャッシュメモリの構成を概念を示す説明図である。

10 第4図はLRU情報フィールドを構成するビットの意味を示す説明図である。

第5図はキャッシュメモリの詳細を例示するブロック図である。

第6図はページ情報エントリの一例を示す説明図である。

15 第7図は内部メモリユニットで実現されるアドレス変換及びキャッシュ制御機能の一例を概略的に示すフローチャートである。

第8図はキャッシュラインのリブレース位置を決定する制御内容を概念的に示す説明図である。

第9図は第7図のステップS6におけるウェイ選択処理の詳細を例示するフローチャートである。

20 第10図はデータプロセッサに適用されるプログラム開発環境の説明図である。

第11図はウェイ指定指示文の使用例を示す説明図である。

第12図は実行可能形式プログラムが含むセクション情報の一例を示す説明図である。

25 第13図は開発されたプログラムを実行するパーソナルコンピュータ又はエンジニアリングワークステーションなどのコンピュータが例

示する概念図である。

発明を実施するための最良の形態

《データプロセッサ》

- 5 第1図には本発明に係るデータ処理システムの概略が示される。同図においてデータ処理システムは代表的に示されたデータプロセッサ1とメインメモリ2を有する。データプロセッサ1は代表的に示された演算制御部3及び内部メモリユニット4を有する。内部メモリユニット4はキャッシュメモリ(CACHE)5、アドレス変換バッファ(TLB)6、及びメモリ管理部(MMU)7を有する。キャッシュメモリ5は命令用とデータ用に分離した分離型キャッシュ構成とされる。命令用キャッシュメモリとデータ用キャッシュメモリはともにウェイを4つもつ4ウェイ構成とされる。命令列およびデータで構成されるプログラムは、
- 10 メインメモリ2に配置されており、プログラム実行すなわち個々の命令の実行は、データプロセッサ1内部の演算制御部3で行なわれる。プログラム実行に必要な命令およびデータの一部は内部メモリユニット4のキャッシュメモリ5に必要に応じて蓄えられていく。プログラム実行が進むにつれて、必要な命令またはデータがキャッシュメモリ5に存在しない場合、すなわちキャッシュミスが発生すると、メモリ管理部7は
- 15 外部バスE X - b u sを通してメインメモリ2から必要な命令またはデータをキャッシュメモリ5に蓄える。演算制御部3は必要な命令又はデータが利用可能になり次第、プログラム実行を再開する。
- 20

- データプロセッサが実行するプログラムの仮想アドレス空間のページ情報は、メインメモリ2のアドレス変換テーブル15にページ情報エントリ40として保持されている。TLB6はページ情報エントリ40
- 25 を高速にアクセスするためのバッファである。このページ情報エントリ

40には仮想ページと物理ページの対応と共にキャッシュメモリ5の
ウェイを指定するウェイ指定フィールド42を有する。キャッシュメモ
リ5は、キャッシュミスに応答する記憶情報のリブレースにおいて、ア
クセスアドレスのページに対応するウェイ指定フィールド42の情報
5 によって指定されるウェイの中からキャッシュフィル対象ウェイを選
択する。その詳細は後述する。

第2図にはデータプロセッサの詳細が例示される。同図に示されるデ
ータプロセッサ1は、例えば公知の半導体集積回路製造技術によって単
結晶シリコンのような1個の半導体基板に形成される。このデータプロ
セッサ1は、特に制限されないが、システムバスS-bus、キャッシュ
10 バスC-bus、及びペリフェラルバスP-busから成る3バス構
成を有する。夫々のバスS-bus, C-bus, P-busはデータ、
アドレス、制御信号の各信号線群を備えている。システムバスS-bu
sには中央処理装置(CPU)3A、乗算器(MLT)3B、キャッシュ
15 ユメモリ(CACHE)5、アドレス変換バッファ(LTB)6、メモ
リ管理部(MMU)7が結合される。CPU3A及びMLT3Bは前記
演算制御部3の一例である。前記キャッシュメモリ5及びメモリ管理部
7は他方においてキャッシュバスC-busに結合され、当該キャッシ
ュバスC-busにはバスステートコントローラ(BSC)11が接続
20 される。バスステートコントローラ11は入出力回路(EXIF)12
を介して外部バスEX-busに結合され、当該外部バスEX-bus
に結合された前記メインメモリ(MMR Y)2等とキャッシュバスC-
busとのインタフェース制御を行う。バスステートコントローラ11
はペリフェラルバスP-busに接続された周辺モジュール(PMDU
25 L)13に対するバスアクセス制御を行う。周辺モジュール13は例え
ばタイマ、シリアルコミュニケーションインタフェース等である。

データプロセッサ 1 は論理アドレス空間を論理ページと呼ばれる単位に分割し、そのページ単位に物理アドレスへのアドレス変換を行うための仮想記憶をサポートする。前記メモリ管理部 7 はアドレス変換バッファ 6 の管理と共にアドレス変換に伴う制御を行う。アドレス変換バッファ 6 は論理ページ番号と物理ページ番号とに関する変換対などを T L B エントリとして格納する連想メモリとして構成され、メモリ管理部 7 は中央処理装置 3 A が出力する論理アドレスをアドレス変換バッファ 6 などを用いて物理アドレスに変換する。T L B ミスの場合にはその論理アドレスに対応される T L B エントリはメモリ管理部 7 を介して外部メモリ 2 上のアドレス変換テーブル（ページテーブル）15 から読み込まれる。前記アドレス変換バッファ 6 はマルチウェイ例えば 4 ウェイ・セットアソシアティブ方式または完全連想方式のキャッシュメモリによって構成される。T L B ミスなどのアドレス変換に係る各種例外が発生すると、メモリ管理部 7 はその例外要因を要因レジスタ（図示せず）にセットし、且つ、T L B ミスなどのアドレス変換に係る例外発生 15 の通知信号（図示せず）を中央処理装置 3 A に送る。中央処理装置 3 A は、要因レジスタにセットされた要因を用いて、或いはそれを用いずにハードウェアで直接、所定の例外処理に分岐される。

中央処理装置 3 A は、例えば 4 ギガバイトの論理アドレス空間をサポートするために 32 ビットのアドレスを利用する。中央処理装置 3 A は、特に図示はしないが、汎用レジスタ、演算器、プログラムカウンタなどの制御用レジスタ群、そして命令のフェッチや解読並びに命令実行手順を制御したり演算制御を行う命令制御部を有する。この中央処理装置 3 A はメインメモリ 2 若しくはキャッシュメモリ 5 から命令をフェッチ 20 し、その命令を命令制御部で解読することにより、当該命令記述に応じたデータ処理を行う。

キャッシュメモリ5はマルチウェイ形式を有し、例えば、4ウェイ・セットアソシアティブ形式の連想メモリ部としてのキャッシュメモリ部とその制御部を備える。キャッシュメモリに対するインデックスは論理アドレスの一部を用いて行われ、キャッシュエントリのタグ部には物理アドレスが保有され、インデックスされたタグ部はその論理アドレスがアドレス変換バッファ6を用いて変換された物理アドレスと比較され、その比較結果に応じてキャッシュミス/ヒットを判定する。キャッシュミスの場合に当該キャッシュミスに係るデータ又は命令はメインメモリ2から読み込まれ、読み込まれたデータ又は命令は新たなキャッシュエントリとしてキャッシュメモリ5に格納される。

第3図にはキャッシュメモリの構成概念を示す。キャッシュメモリ5はキャッシュセットSET1～SETnの集まりであり、各キャッシュセットは4ウェイWAY1～WAY4を有し、ウェイ毎に有効フラグV、タグ情報T、データラインL等を持つ。各セットにおけるウェイ毎の記憶エリアをキャッシュラインと称する。前記データラインLはメインメモリ2においてアドレス上連続するデータの予め決められたサイズ分のコピーである。本実施例ではデータラインLのサイズを32バイトとする。アクセスアドレス20はタグアドレス21、インデックスアドレス22、及びラインオフセット23として把握することができる。前記タグ情報Tはこれと対を成すデータラインLの識別情報であり、アドレス情報に含まれるタグアドレス21と比較されて参照される。ラインオフセット23は参照したいデータのデータラインL上での位置を示すオフセットアドレスであり、データラインサイズが32バイトであればアドレス情報20の下位5ビットがラインオフセット23になる。有効フラグVはそれと対を成すデータラインLが有効であるか否かを示す。さらに各セットSET1～SETnは夫々LRU情報フィールド24

を持つ。このLRU情報フィールド24は、対応するセットのどのウェイを新たなデータの充填先に選ぶかを決定する際に参照されるフィールドであり、キャッシュミスに応答する記憶情報のリプレースにおいて、どのウェイをキャッシュフィルの対象にするかを選ぶのに用いられる

5 情報である。LRU情報フィールド24は、4ウェイの場合、6ビットによって構成される。このLRU情報フィールド24を構成するビットの意味は第4図に例示される。LRU情報フィールド24の値は、リプレースの候補とされるウェイの中で最も最近参照されていないウェイを選択するための情報として利用される。

- 10 第3図に示されるアドレス情報20のインデックスアドレス22は、キャッシュセットの集まりから、参照したいデータが格納されているはずのキャッシュセットを特定する情報である。ここでは256個のキャッシュセットがあり、この場合、インデックスアドレス22は8ビットとされる。アドレス情報20はTLB6のページ情報から構成される物理
- 15 アドレスとする。ただし、インデックスアドレス22とラインオフセット23は仮想アドレスの対応するビットフィールドと一致する。キャッシュメモリの基本的な構成はデータキャッシュメモリも命令キャッシュメモリも変わらない。

- 第5図にはキャッシュメモリの詳細が例示される。キャッシュメモリ
- 20 5は、4個のウェイWAY1～WAY4を構成するメモリアレイ31M～34M、インデックスアドレス22に基づいて夫々のメモリアレイ31M～34Mからキャッシュラインを選択するインデックスアドレスデコーダ31D～34D、比較回路CMP1～CMP4、リードデータセクタRSEL、ライトデータセクタWSEL、及び制御部CNT
- 25 を備える。インデックスアドレスデコーダ31D～34Dはインデックスアドレス22に基づいて各ウェイWAY1～WAY4から夫々キャ

ッシュラインを選択し、これにより、インデックスアドレス 2 2 に対応する一つのキャッシュセットが選択される。比較回路 CMP 1 ~ CMP 4 は、選択されたキャッシュセットの対応するウェイから読み出される有効フラグ V とタグ情報 T を入力し、有効フラグ V が対応データライン L の有効性を意味し、且つタグ情報 T とタグアドレス 2 1 が一致する（キャッシュヒット）か否かを判定し、ウェイ毎の判定結果をヒット信号 3 0 にて制御部 CNT に与える。

制御部 CNT は、ヒット信号 3 0 によりキャッシュヒットを判定すると、リードアクセスに対してはリードセクタ RSEL に、キャッシュヒットに係るウェイから読み出されたデータライン L のデータを選択させ、その中からラインオフセット 2 3 で指定されるデータを切出して出力させる。ライトアクセスのキャッシュヒットに対してはライトセクタ RSEL に、インデックスアドレス 2 2 によって選択されるキャッシュセットの中のキャッシュヒットに係るウェイのデータライン L の中のラインオフセット 2 3 で指定される位置に向けてライトデータを出力させる。

特に図示は省略するが、制御部 CNT はインデックスされたキャッシュセットに含まれる各キャッシュラインの有効フラグ V を入力する。制御部 CNT は、ヒット信号 3 0 によりキャッシュミスと判定したとき、インデックスアドレス 2 2 によって選択されるキャッシュセットの中で、有効フラグ V によって無効が示されているウェイがあれば、そのデータライン L にキャッシュミスに係るデータがフィル（充填）される。一方、インデックスアドレス 2 2 によって選択されるキャッシュセットの中で、有効フラグ V によって無効が示されるウェイが無ければ、当該キャッシュセットの中から LRU 情報フィールド 2 4 の LRU 情報及びウェイ指定フィールド 4 2 の情報を用いて一つのウェイを選択し、選

択したウェイに対してキャッシュデータエントリのリプレースを行う。即ち、リプレースすべきウェイの元のキャッシュデータエントリの内容がメインメモリに反映されていなければ、そのデータをメインメモリに退避し、その後で、キャッシュミスに係るデータがフィルされる。

- 5 図6にはページ情報エントリの一例が示される。ページ情報エントリ40は仮想アドレス空間を構成する各ページ毎に存在され、これらの集合によってページ情報テーブル15が形成される。このページ情報テーブル15の一部のエントリがTLB6に格納される。

- 10 仮想ページ番号43はプログラムの仮想空間における仮想ページ番号であり、TLB6を検索するキーとなる情報である。物理ページ番号44は仮想ページがメインメモリのどのページに対応するかを示す情報であり、メインメモリにアクセスするときはこの情報と仮想アドレスから物理アドレスを生成する。ページ情報フィールド41はこのページの属性情報をもつ。例えば、このページが属するプログラムを識別する
- 15 プロセス情報や、このページへの書き込みの可否を示す情報、キャッシュへの格納の可否を示す情報、このエントリの有効性を示す情報などが格納されている。ウェイ指定フィールド42はこのページに属するデータまたは命令がマルチウェイ方式キャッシュのどのウェイに位置すべきかを示す情報である。この例では、ウェイ指定フィールド42は4ビットのウェイ指定ビット（ウェイ指定情報）W1～W4を有する。例えば
- 20 ウェイ1の指定ビットW1は、“1”であれば当該ページに属する命令またはデータがキャッシュメモリ上で充填（キャッシュフィル）されるウェイ候補の中にウェイWAY1を含むことを意味し、“0”であれば当該ページに属する命令またはデータがキャッシュフィルされるウェイ候補にはウェイWAY1を選ばないことを意味する。その他のウェイ指定ビットW2～W3も同様に、それぞれウェイWAY2、WAY3、
- 25

- WAY 4 が充填先候補か否かを指定する意味を持つ。例えば、ウェイ指定フィールド 4 2 の値が “W 1 W 2 W 3 W 4 = 1 0 1 1” なら、これはウェイ WAY 1、ウェイ WAY 3、ウェイ WAY 4 を充填先候補とし、ウェイ WAY 2 は充填先としないことを意味する。また、ウェイ指定フィールド 4 2 の値が “1 1 1 1” なら、4 つのウェイ WAY 1 ~ WAY 4 すべてを充填先候補とする。充填候補が複数の場合、ここでは、当該複数の充填候補の中で L R U 論理に従った一つのウェイが選択されることになる。尚、ページ情報フィールド 4 1 によりキャッシュメモリに配置可能と指定されているページのウェイ指定情報 4 2 が全ビット
- 5 “0” の場合は、当該ウェイ指定情報 4 2 が読み出された段階で実行時エラーとし、ユーザ規定の例外処理の対象にすることが可能である。

- 第 7 図には内部メモリユニット 4 で実現されるアドレス変換及びキャッシュ制御機能の一例が概略的に示される。ここでは、一例として、仮想アドレス a に位置するデータ x をデータプロセッサ 1 上のある汎用レジスタ r にもってくるロード命令を実行する場合を考える。ステップ S 1 で、仮想アドレスから該当するページ情報を T L B 6 から引き出す。ステップ S 2 で仮想アドレス a に対応するページ情報エントリが T L B 6 に存在するかを判定し、存在しない場合、ステップ S 3 の T L B ミス例外処理を行う。これは、メインメモリ 2 に格納されているアドレス変換テーブル 1 5 またはプログラムの実行可能ファイルを参照して該当するページ情報エントリを T L B 6 に登録する処理である。T L B 6 にページ情報エントリが存在する場合、該当ページ情報エントリ 4 0 からページ情報フィールド 4 1、ウェイ指定フィールド 4 2 及び物理ページ番号 4 4 を取り出す。
- 15
- 20
- 25
- ステップ S 4 では仮想アドレス a に対応する情報をキャッシュメモリ 5 から取り出す。取り出す情報はインデックスアドレス 2 2 で示され

るキャッシュセットを構成する各ウェイ毎の有効フラグV、タグ情報T及びLRU情報24である。

5 ステップS5はステップS4で取り出された情報から該当するデータがキャッシュメモリ5に存在するか否かを判定する。その判定は、タグアドレス21と各ウェイのタグTが一致し、且つ対応する有効フラグVがキャッシュラインの有効性を示すとき、該当するデータがキャッシュメモリ5にある（キャッシュヒット）と判定する。

10 例えば必要なデータ又はプログラムがキャッシュメモリ5に格納されているなら、ステップS9によりそのデータ又は命令を演算部制御部3に供給する。ステップS5においてキャッシュミスであると判定されたときは、ステップS6で仮想アドレスaに対応するデータxを充填すべきキャッシュメモリ5上のウェイを決定し、ステップS7で充填すべきデータをメモリアクセスし、ステップS8でそのデータを所定のウェイのキャッシュラインに充填する。

15 第8図はキャッシュラインのリブレース位置を決定する制御内容を概念的に示す。第7図と同様にアドレスxのインデックスアドレス22に対応されるセットのセット情報（V，T，LRU）50はセット選択手段51により選択され、ウェイ選択手段52に入力される。そして、前記アドレスxに対応されるページ情報エントリ40内のウェイ指定
20 フィールド42の情報W1～W4もウェイ選択手段52に入力される。ウェイ選択手段52は、ウェイ指定フィールド42で指定されるウェイ候補に対するLRU(Leas Recently Used)アルゴリズムの適用により充填先ウェイを決定する。アドレスxに対応してメインメモリ2から読み出された充填されるべきキャッシュフィールドデータ（充填データ）53の
25 キャッシュメモリ上の充填先は、セット選択手段51によりインデックスアドレスで一意的に指定されたセットに含まれる4つのキャッシュ

ラインのうち、ウェイ選択手段52で決定されたウェイに対応するキャッシュラインとなる。第8図のセット選択手段51は第5図の各メモリマットのインデックスアドレスデコーダ31D~34Dによって実現され、ウェイ選択手段52は第5図のコントローラCNT及びライトセレクトタWSELによって実現される。

- 第9図には第7図のステップS6におけるウェイ選択処理の詳細が例示される。キャッシュミスが発生すると、先ずステップS10で、先にインデックスされている充填先キャッシュセットにおける各キャッシュラインのセット情報50のうち、ウェイ指定フィールド42で充填候補に指定されているウェイの有効フラグVをチェックする。夫々のチェック結果に基づいて、ステップS11では、全てのキャッシュラインの有効フラグVが有効か否かを判定する。無効を意味するキャッシュラインがあれば、すなわち、インデックスされたセットに無効なキャッシュラインがあれば、ステップS12に進み、そのような無効なキャッシュラインの中から充填先キャッシュラインを1つ選択する。ここでは、そのようなウェイのうち最小のウェイ番号をもつものを充填先ウェイとする。もし、ステップS11で、対応するウェイがすべて有効であれば、それらのどれかに対応するキャッシュラインを新しいデータで置き換える必要があり、次のステップS13に進む。
- 20 ステップS13では当該キャッシュセットのLRU情報を読み出す。これは第3図のLRU情報フィールド24を構成する6ビットのうち、ウェイ指定フィールド42の情報で指定された充填先ウェイ候補間のアクセス時期の前後関係に関するビットを読み出すことである。例えば、ウェイ指定フィールド42の値が“0111”であれば、充填先ウェイ候補はウェイ2(WAY2)、ウェイ3(WAY3)、ウェイ4(WAY4)となり、これらのアクセス時期関係を示すのは第4図より、LR

U情報フィールドの第4ビット、第5ビット、第6ビットとなる。いま、LRU情報フィールド24が“000000”だったとする。第4図より、 $WAY1 < WAY2 < WAY3 < WAY4$ という関係が判明する。ここで大小関係 $WAY2 < WAY3$ は $WAY2$ のアクセス時期が $WAY3$ のアクセス時期より古いことを表すとする。

ステップS14で充填先ウェイを決定する。いま、LRU情報フィールドが“000000”だったとする。一方ウェイ指定フィールドが“0111”なので $WAY2$ 、 $WAY3$ 、 $WAY4$ のアクセス時期関係を表すLRU情報フィールド内の第4ビット、第5ビット、第6ビットをみるとそれぞれ、“0”、“0”、“0”なので、第4図より、 $WAY2 < WAY3 < WAY4$ という関係が判明する。従って、この例では、置き換え対象となるウェイは最もアクセスされた時期が古いウェイ $WAY2$ となる。

ステップS15では、ステップS14で選択されたウェイに対応するラインLのデータをメインメモリ2に書き戻す。

ステップS16ではLRU情報フィールドを更新する。今の例でいえば、ウェイ $WAY2$ が最も最近アクセスされたという情報に更新する。これはウェイ $WAY2$ と他のウェイのアクセス時期の関係を示す3つのビットを更新すればよく、第4図より、更新対象となるビットはLRU情報フィールドの第1ビット、第4ビット、及び第5ビットとなり、これら3つのビットをそれぞれ“0”、“1”、“1”にすればよい。この結果、LRU情報フィールドは“000110”となり、これは「 $WAY1 < \text{ウェイ}WAY3 < WAY4 < WAY2$ 」というアクセス時期関係を表す。

ステップS17では、ステップS14で選択されたウェイに対応するラインを新たなデータで充填し、有効フラグV、タグTを更新する。キ

キャッシュメモリに供給されたデータは同時に前記演算制御部 3 に供給され、キャッシュミスでストールしていた演算が再開される。

上記キャッシュメモリ 5 に関するウェイ指定フィールドの活用とライン充填手順は、データ用キャッシュメモリだけでなく、命令用キャッシュメモリ、命令及びデータ混在型のユニファイドキャッシュメモリにも適用できる。例えば、命令キャッシュメモリに適用する場合、プログラムの実行に伴ってアドレス x の命令が必要になったとき、当該アドレスのデータが命令用キャッシュメモリにないとき、当該アドレスに対応するページ情報エントリのウェイ指定フィールドと、命令用キャッシュ内 LRU 情報フィールドから、新しい命令の充填先ラインを決定すればよく、その手続は前述と同様の充填先ライン決定手順を用いればよい。ユニファイドキャッシュメモリの場合も同様である。さらに、データプロセッサ 1 の外部に配置したキャッシュメモリに対しても上述のキャッシュフィル対象ウェイ選択制御手順を適用することができる。

また、キャッシュメモリのウェイ数は 4 ウェイ以外であっても、ウェイ指定フィールドのビット数、LRU 情報フィールドのビット数を拡張し、拡張に合わせてウェイ選択方式を変更すれば、上述とほぼ同様の手順で充填先となるキャッシュラインを決定することが可能である。

《プログラム開発方法》

次に、プログラム開発方法の観点より前記キャッシュフィル対象ウェイ選択制御について説明する。ここでは、プログラミング言語として C 言語を例に説明する。

まず、第 10 図に基づいてデータプロセッサのプログラムの開発環境について概略を説明する。プログラマは、各種エディタなどを用いて、例えば C 言語でプログラム（C 言語ソースプログラム）を作成する。これは通常、複数のモジュールに分割して作成される。C コンパイラ 60

は、プログラムの作成したそれぞれのC言語ソースプログラムを入力し、オブジェクトモジュールを出力する。リンカージェディタ61は、上記Cコンパイラ60で生成した、複数のオブジェクトモジュール61を入力して、各モジュールの外部参照や相対アドレスなどの解決を行い、1つのプログラムに結合して、ロードモジュール63を出力する。ロードモジュール63は、シミュレータ/デバッガ64に入力して、演算器シミュレータとともに、パーソナルコンピュータなどのシステム開発装置上で、データプロセッサ1の動作をシミュレーションし、実行結果を表示し、プログラムの解析や評価を行なうことができる。また、ロードモジュール63をエミュレータ65にダウンロードして、実際の応用システム上などで動作する、いわゆるインサーキットエミュレーションを行ない、システム全体としての、実動作の解析や評価を行なうことができる。

C言語では、言語仕様以外の特別な情報をコンパイラ60に伝える手段として指示文と呼ばれる特殊な文でその言語処理系特有の処置をコンパイラに指示することができる。この指示文はプリAGMAとも呼ばれ、例えば一般に「`#pragma information_x ("function1")`」などと記述される。この例では「`information_x`」という名前がこの指示文の機能を代表するものであり、「`function1`」というものがその引数である。たとえば、これは「`function1`」というシンボル名をもつ関数に対して「`information_x`」で表現される処置を施すようにコンパイラに指示するものである。

ここでは、プログラムを構成するコードおよびデータがプログラム実行時にキャッシュメモリのどのウェイに割り付けられるかを指示するために前記指示文を利用する。要するに、プログラマはプログラムを構成するコードおよびデータをプログラム実行時にキャッシュメモリの

どのウェイに割り付けられるかを任意に指示することができる。

第11図はウェイ指定指示文の使用例が示される。第11図の(a)に示す構文をもつ指示文によりプログラマはプログラムを構成する命令やデータをセクションと呼ばれる、仮想アドレス空間上で連続する領域に集めるようコンパイラに指示できる。ここでは名前「section」で表される指示文は、「シンボル名」、「セクション名」、「ページサイズ」、「ウェイ指定」という4つの引数をもつ。「シンボル名」にはプログラム内の変数や関数名のリストを記述する。「セクション名」は「シンボル名」で記述された変数や関数をまとめて配置する仮想
5 アドレス空間上で連続する領域の名前を記述する。「ページサイズ」には、そのセクションの基本ページサイズを記述する。そして、「ウェイ指定」にはそのセクションに属するデータや命令がどのキャッシュメモリのどのウェイに割り付けられるかを指示する情報が記述される。

第11図の(b)に二つの正方行列を配列Aと配列Bで表し、その二
15 つの行列積を計算するコード例を示す。配列Aおよび配列Bのサイズがそれぞれデータ用キャッシュサイズを上回るとき、このループが実行されていくにつれ、データ用のキャッシュメモリ上に既に存在するデータであって後の処理で再利用されるデータがあっても、従来、それらは全てデータ用のキャッシュメモリから追出される可能性がある。また、配
20 列Aおよび配列Bがデータ用のキャッシュメモリ上でキャッシュラインを競合する可能性もある。そこで、前述のウェイ指定制御を適用できるように、配列Aおよび配列Bのデータがキャッシュメモリ上の特定ウェイに割り付けられるようにするために、例えば、第11図の(d)のように指示文を指定する。この例でいえば、変数名Aのデータは「dataA」というセクションに配置し、このセクションの基本ページサイズは4キロバイトであり、このセクションに属するデータはウェイ指定
25

「0100」であり、これはウェイ2をdataAセクションの割り付け先ウェイと指示している。同様に変数名Bのデータは「dataB」というセクションに配置し、このセクションの基本ページサイズは4キロバイトであり、このセクションに属するデータはウェイ指定「0010」とであり、これはウェイ3をdataBセクションの割り付け先ウェイと指示している。これにより、配列Aおよび配列B以外のデータで、後の処理でも再利用されるものをウェイ1に割り付けるように指示しておけばデータ用キャッシュメモリのウェイ1は大規模な配列データによりメインメモリに追いやられることがなくなり、これら再利用可能なデータアクセスでのキャッシュミスを予防することができる。また、配列Aと配列Bの割り付けられるウェイが異なるので、配列Aのデータと配列Bのデータ間でのラインの取り合い（ライン競合という）も予防できる。

第11図の(c)はファイル「code1.c」の内部で「myfunc」という関数を定義している。「myfunc」はファイル「libraryA.c」に属する関数「libfunc」を呼び出している。ここで、ファイル「libraryA.c」内の関数が他の関数からも頻繁にアクセスされるため、できるだけ命令用キャッシュメモリに留めておきたいと仮定する。そのため、第11図の(e)に示すような指示文を記述する。まず、ファイル「code1.c」内の関数は「mycode」セクションにまとめ、このセクションが割り付けられる命令用キャッシュ上のウェイはウェイ指定「1100」によりウェイ1およびウェイ2とする。一方、ファイル「libcodeA」内では関数「libfunc」を含む関数群をセクション「libcodeA」にまとめ、このこのセクションが割り付けられる命令用キャッシュ上のウェイはウェイ指定「0011」によりウェイ3およびウェイ4とする。これ

によりファイル「libcodeA.c」に記述される関数は命令用キャッシュ上で関数「myfunc」などとのライン競合により追い出されることが少なくなるようにできる。

特にウェイ指定されないデータまたはプログラムコードについては
5 それらが属するデフォルトのセクションのウェイ指定情報は、「1111」すなわち、キャッシュのすべてのウェイを活用するとコンパイラは解釈する。

一般に個別のソースファイルを個別にコンパイルするとオブジェクトとよばれるファイルが作成される。実行可能形式のプログラムを作成
10 するには、これらオブジェクトファイルをまとめて1つの仮想アドレス空間にする必要がある。この操作はリンクと呼ばれ一般にリンカまたはリンケージエディタがこの操作を行う。第11図の(a)のウェイ指定指示文で指定された情報に基づいて、リンカは実行可能形式のプログラムのなかのセクション情報にウェイ指定情報を埋め込む。

15 第12図には実行可能形式プログラムが含むセクション情報の一例を示す。リンカは実行可能形式プログラムの各セクションごとに第12図に示すセクション情報を作成する。このセクション情報の「スタートアドレス」は仮想アドレス空間での当該セクションの開始アドレスを意味する。セクション情報の「セクションファイルオフセット」は実行可能
20 プログラムを格納しているファイル内での当該セクションの配置位置、「セクションサイズ」は当該セクションのサイズ、「基本ページサイズ」は当該セクションの基本ページサイズである。そして「ウェイ指定情報」は、当該セクションに属するデータまたは命令が実行される時にそれらをキャッシュメモリのどのウェイに割り付けられるかを指示
25 する情報であり、第11図の(a)で指定されたウェイ指定情報が反映される。

データプロセッサ 1 のオペレーティングシステム (OS) は実行形式プログラムを実行するときまず、プログラムの開始命令を含むページを TLB 6 に登録する。このとき実行可能形式プログラムのセクション情報内のウェイ指定情報を参照し、ウェイ指定フィールド 42 に書きこむ。

- 5 そしてプログラムの実行に伴い初めて参照されるページに対するページ情報エントリ 40 の登録の際も同様に実行形式プログラムのセクション情報を参照し対応するページ情報エントリのウェイ指定フィールド 42 を指定する。

- 10 このように、プログラミング言語の指示文により、プログラム実行時に前記所定の情報をキャッシュメモリのどのウェイに割り付けるかを指示するようにすることにより、ソースレベルのプログラム作成段階からプログラマはキャッシュのウェイの運用方法を指示する簡易な手段を活用して、プログラム実行時のキャッシュミスによる性能劣化を予防できる。さらに、データプロセッサのプログラム開発において、データ
15 やプログラムの利用形態若しくは性質に即応させてそれら情報に対するウェイへの割り当てを任意に決定できる、というキャッシュ制御の自由度を簡単に得ることができる。

- 20 前述の開発環境により、プログラムのアドレス空間を管理する仮想記憶とマルチウェイ形式のキャッシュメモリとを用いてデータ処理を行うためのプログラムが開発され、開発されたプログラムは、コンピュータ読取り可能に記録媒体に記録して提供することが可能である。例えば、第 13 図には開発されたプログラムを実行するパーソナルコンピュータ又はエンジニアリングワークステーションなどのコンピュータ 70 が例示される。前記開発環境で開発されたプログラムは単数又は複数枚
25 のフロッピディスク (FD) 71 や CD-ROM ディスク 72 等の記憶媒体に原始的に記録されて提供され、それら記録媒体に記録されたプロ

グラムは計算機内部のハードディスク装置 7 3 などの記録媒体にインストールされて保持される。或いは、そのようなプログラムは、公衆回線網 7 4 などを介してダウンロードされ、前記ハードディスク装置 7 3 に記録されている。

- 5 このとき、そのプログラムは、アドレス空間上における前記プログラムの配置を規定する管理情報としての第 1 2 図のセクション情報を有し、この管理情報の中に、プログラムを構成する所定の情報をキャッシュメモリのどのウェイに割り付けるかを指示するウェイ指定情報が含まれる。コンピュータ 7 0 がそのプログラムを記録媒体 7 3、7 2、7 1 から読み取って実行することにより、前述のキャッシュメモリ 5 における記憶情報のリプレースに際して、時間的にも空間的にも再利用性の比較的高い情報が再利用性の比較的低い情報によって追い出され事態の防止を容易に実現できる。また、前記記録媒体 7 3、7 2、7 1 に記録されたプログラムが携帯情報端末装置 (PDA) 7 5 の動作プログラムである場合、コンピュータ 7 0 にシリアルインタフェースなどを介して携帯情報端末装置 7 5 を接続し、コンピュータ 7 0 が前記記録媒体 7 3、7 2、7 1 からプログラムを読み取って携帯情報端末装置 7 5 にダウンロードしてもよい。この場合には、携帯情報端末装置 7 5 に内蔵されたキャッシュメモリ 5 における記憶情報のリプレースに際して、時間的にも空間的にも再利用性の比較的高い情報が再利用性の比較的低い情報によって追い出され事態の防止を容易に実現できる。

以上本発明者によってなされた発明を実施例に基づいて具体的に説明したが本発明はそれに限定されるものではなく、その要旨を逸脱しない範囲において種々変更可能である。

- 25 例えば、命令とデータに共用されるキャッシュメモリの場合も第 6 図のウェイ指定フィールド 4 2 を活用したウェイの運用が可能である。メ

メモリ管理部 7 がそのようなデータ・プログラム混在型のキャッシュ管理
に対応すればよい。また、命令・データ共有キャッシュメモリを採用す
るデータプロセッサ、例えば、チップ面積を抑えたい組み込み用途のプ
ロセッサなどでも、プログラマはキャッシュメモリのウェイの運用方法
5 を指示する前記簡易な手段を活用して、プログラム実行時のキャッシュ
ミスによる性能劣化を予防できる。例えば、命令・データ共有キャッシ
ュでもプログラムの特性に応じて命令用あるいはデータ用ウェイなど
を指定できる。

キャッシュメモリはデータプロセッサ内部にある場合に限定されな
10 い。キャッシュメモリがデータプロセッサの外部にある構成でも第 6 図
のウェイ指定フィールド 42 を活用したウェイの運用が可能である。メ
モリ管理部 7 が外部キャッシュメモリ管理に対応すればよい。したがっ
て、データプロセッサの外部にキャッシュをもつ場合、例えば、キャッ
シュ容量を増やした 2 次キャッシュメモリ、3 次キャッシュメモリなど
15 を持つ場合であっても、プログラマはキャッシュメモリのウェイの運用
方法を指示する簡易な手段を活用して、プログラム実行時のキャッシュ
ミスによる性能劣化を予防できる。

以上の説明から明らかなように、本発明によれば、メインメモリから
持ってきたデータまたは命令を格納するキャッシュメモリ内のメモリ
20 領域をそのデータまたは命令列アクセスの特徴に応じて選択的に指定
することが可能となり、再利用性の比較的高いデータ又は命令が再利用
性の比較的低いデータまたは命令により追出される状況を防いでキャ
ッシュミスを低減でき、プログラムの実行性能を向上できる。

さらに、本発明によれば、従来のメモリ管理装置の機構を大幅に変更
25 することなく、プログラミング上の自由度が高められプログラムの実行
性能を向上するようなキャッシュメモリを意識したプログラミングが

容易となる。

産業上の利用可能性

- 本発明は、機器組み込み制御用途のチップ面積が制約されるキャッシュメモリ内蔵マイクロコンピュータ等のデータプロセッサ、仮想記憶をサポートし且つキャッシュメモリを用いるデータ処理システム、携帯情報端末装置、パーソナルコンピュータなどに広く適用することができる。
- キャッシュメモリによるキャッシュの対象となるデータは、音声データや画像データに代表されるストリーミングデータであっても、或いはその他のデータであってもよい。音声データと画像データを共にキャッシュの対象にしてもよい。

請 求 の 範 囲

1. 中央処理装置と、前記中央処理装置が実行するプログラムのアドレス空間を構成するページのページ情報を保持するページ情報テーブルと、マルチウェイ形式のキャッシュメモリとを含むデータ処理システムであって、
5 前記ページ情報テーブル内の各ページ情報フィールドは、キャッシュメモリのウェイを指定するウェイ指定フィールドを有し、
 前記キャッシュメモリは、キャッシュミスに応答する記憶情報のリ
10 プレースにおいて、アクセスアドレスのページに対応するウェイ指定
 フィールドによって指定されるウェイの中からキャッシュフィル対
 象ウェイを選択するものであることを特徴とするデータ処理システム。
2. 前記ウェイ指定フィールドは、最大で全てのウェイを指定可能な複数ビットを有して成るものであることを特徴とする請求の範囲第 1
15 項記載のデータ処理システム。
3. 前記キャッシュメモリは、キャッシュミスに応答する記憶情報のリプレースにおいて、前記ウェイ指定フィールドが指定するウェイの中から一つのウェイを選択する選択論理を有するものであることを特徴とする請求の
20 範囲第 2 項記載のデータ処理システム。
4. 前記選択論理は、最も最近参照されていないウェイを選択する L R U 選択論理であることを特徴とする請求の範囲第 3 項記載のデータ処理システム。
5. 前記ページ情報テーブルは、前記中央処理装置によってアクセス可能な
25 メインメモリに形成されるものであることを特徴とする請求の範囲第 1
 項記載のデータ処理システム。

6. 前記ページ情報テーブルに格納されているページ情報の一部をコピーとして保有する連想記憶形式のアドレス変換バッファを更に有して成るものであることを特徴とする請求の範囲第5項記載のデータ処理システム。
- 5 7. 前記キャッシュメモリは、前記中央処理装置及びアドレス変換バッファと同一の半導体基板に形成されて成るものであることを特徴とする請求の範囲第6項記載のデータ処理システム。
8. 前記キャッシュメモリは、前記中央処理装置及びアドレス変換バッファと異なる半導体基板に形成されて成るものであることを特徴とする請求の範囲第6項記載のデータ処理システム。
- 10 9. 中央処理装置と、前記中央処理装置のアドレス空間を管理する為のページ情報エントリの記憶領域を有するアドレス変換バッファと、マルチウェイ形式のキャッシュメモリとを含むデータプロセッサであって、
- 15 前記ページ情報エントリの記憶領域には前記キャッシュメモリのキャッシュミスに応答するキャッシュフィル対象ウェイの候補を指定するウェイ指定フィールドを有し、
- 前記キャッシュメモリは、アクセスアドレスに対応するページ情報エントリに含まれる前記ウェイ指定フィールドの値を入力し、キャッシュミスに
- 20 シュミスに応答する記憶情報のリプレースにおいて、入力した前記ウェイ指定フィールドが指定する候補の中からキャッシュフィル対象ウェイを決定する制御手段を有して成るものであることを特徴とするデータプロセッサ。
10. 前記ウェイ指定フィールドは、キャッシュフィル対象ウェイの候補を複数指定可能な複数ビットを有するものであることを特徴とする請求の範囲第9項記載のデータプロセッサ。
- 25

- 1 1. 前記制御手段は、入力した前記ウェイ指定フィールドが指定する候補の中から一つのウェイを選択する選択論理として、最も最近参照されていないウェイを選択するLRU選択論理を有して成るものであることを特徴とする請求の範囲第10項記載のデータプロセッサ。
- 5 1 2. 1個の半導体基板に前記中央処理装置、アドレス変換バッファ、及びキャッシュメモリが形成されて成るものであることを特徴とする請求の範囲第9項記載のデータプロセッサ。
- 1 3. 前記キャッシュメモリはデータ用と命令用に分離されて成るものであることを特徴とする請求の範囲第9項記載のデータプロセッサ。
- 10 1 4. 前記キャッシュメモリはデータと命令に共用されるものであることを特徴とする請求の範囲第9項記載のデータプロセッサ。
- 1 5. マルチウェイ形式のキャッシュメモリと中央処理装置を備えたデータプロセッサにより実行可能な形式のプログラムの生成を支援するプログラム開発方法であって、
- 15 プログラミング言語の指示文により、プログラムを構成する所定の情報を仮想アドレス空間上の所定の領域にまとめて配置することを指示すると共に、プログラム実行時に前記所定の情報をキャッシュメモリのどのウェイに割り付けるかを指示することを特徴とするプログラム開発方法。
- 20 1 6. プログラムのアドレス空間を管理する仮想記憶とマルチウェイ形式のキャッシュメモリとを用いてデータ処理を行うためのプログラムをコンピュータ読取り可能に記録した記録媒体であって、
- アドレス空間上における前記プログラムの配置を規定する管理情報を有し、この管理情報の中に、プログラムを構成する所定の情報を
- 25 キャッシュメモリのどのウェイに割り付けるかを指示するウェイ指定情報を有するものであることを特徴とするコンピュータ読取り可

能な記録媒体。

17. マルチウェイ形式のキャッシュメモリを有する情報処理装置によりストリーミングデータを処理する処理方法であって、

5 上記ストリーミングデータのキャッシュ対象となるウェイを前記
 キャッシュメモリの一部のウェイの範囲に限定して指定可能とすることを特徴とするストリーミングデータ処理方法。

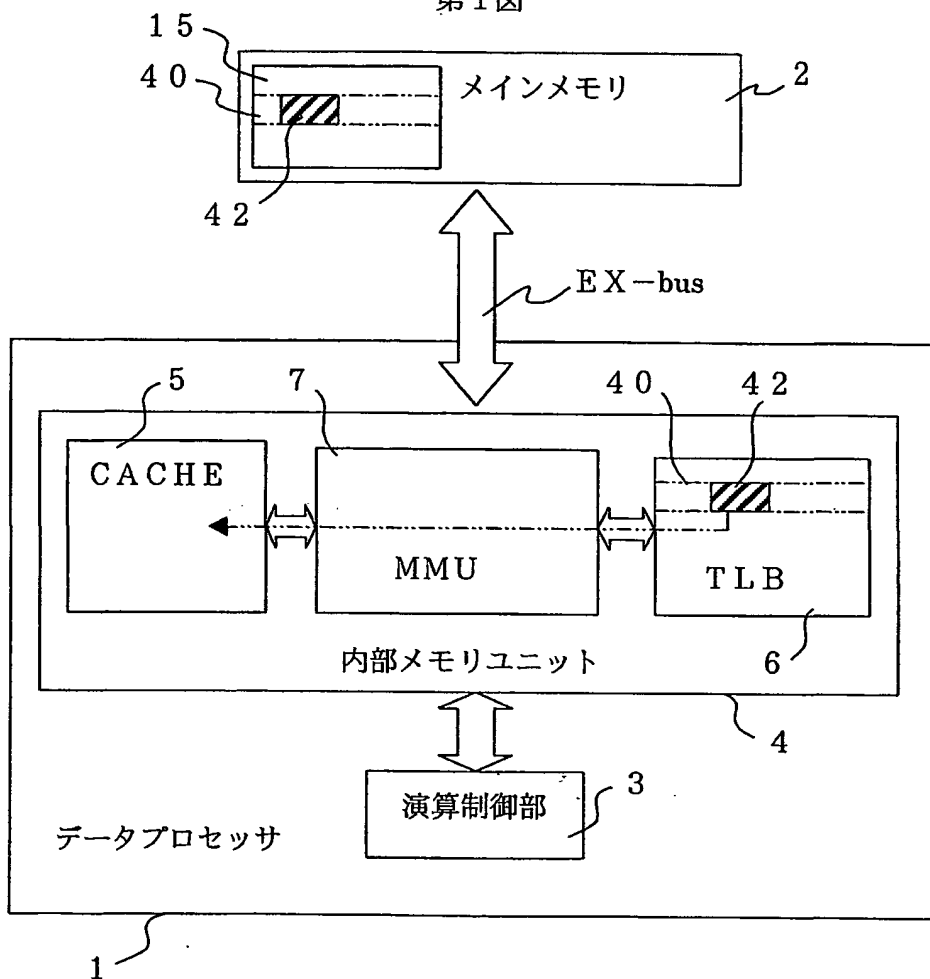
18. 上記ストリーミングデータは音声データであることを特徴とする請求の範囲第17項記載のストリーミングデータ処理方法。

10 19. 上記ストリーミングデータは画像データであることを特徴とする
 請求の範囲第17項記載のストリーミングデータ処理方法。

20. 上記ストリーミングデータは画像データと音声データであることを特徴とする請求の範囲第17項記載のストリーミングデータ処理方法。

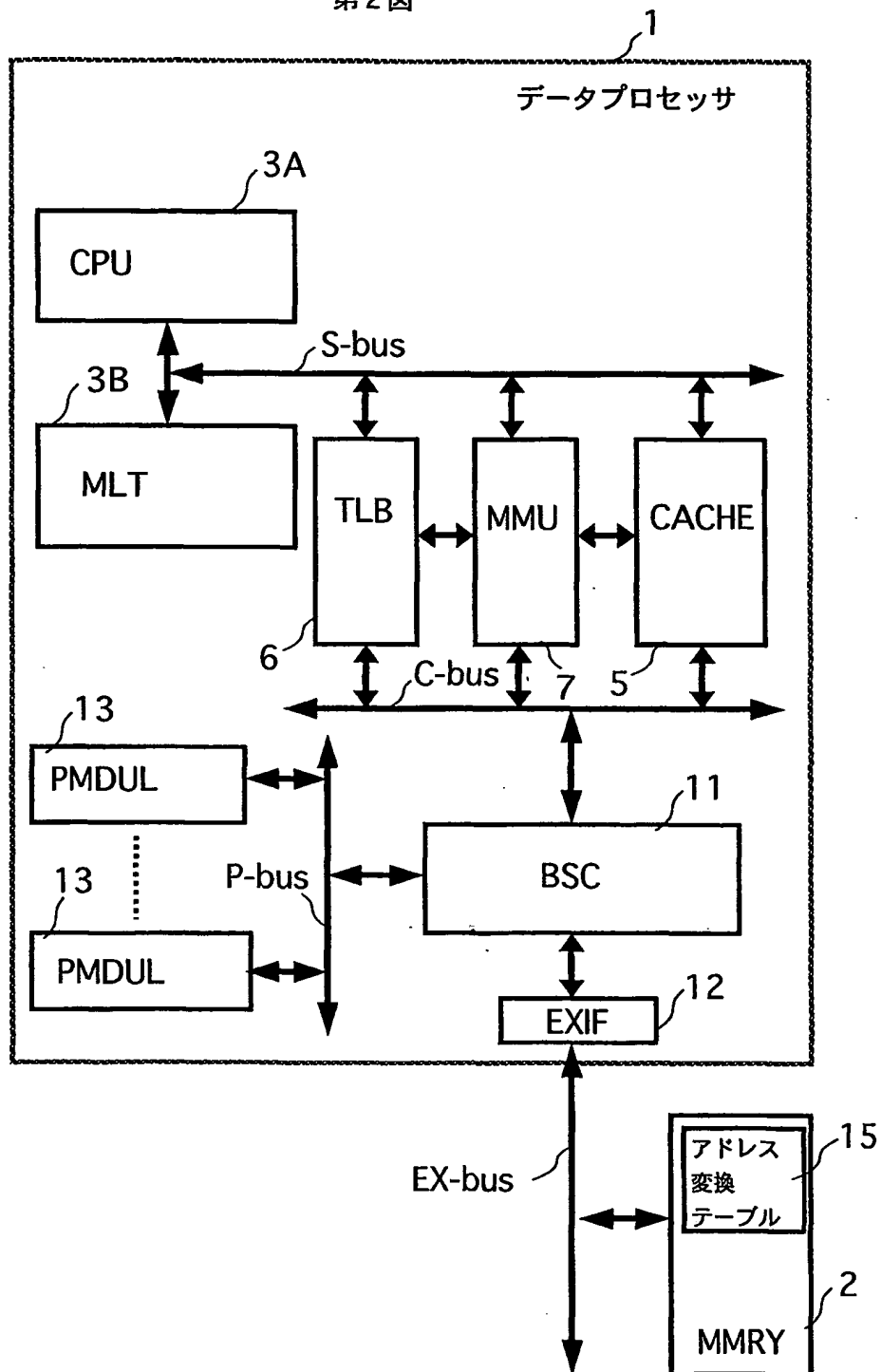
1/12

第1図



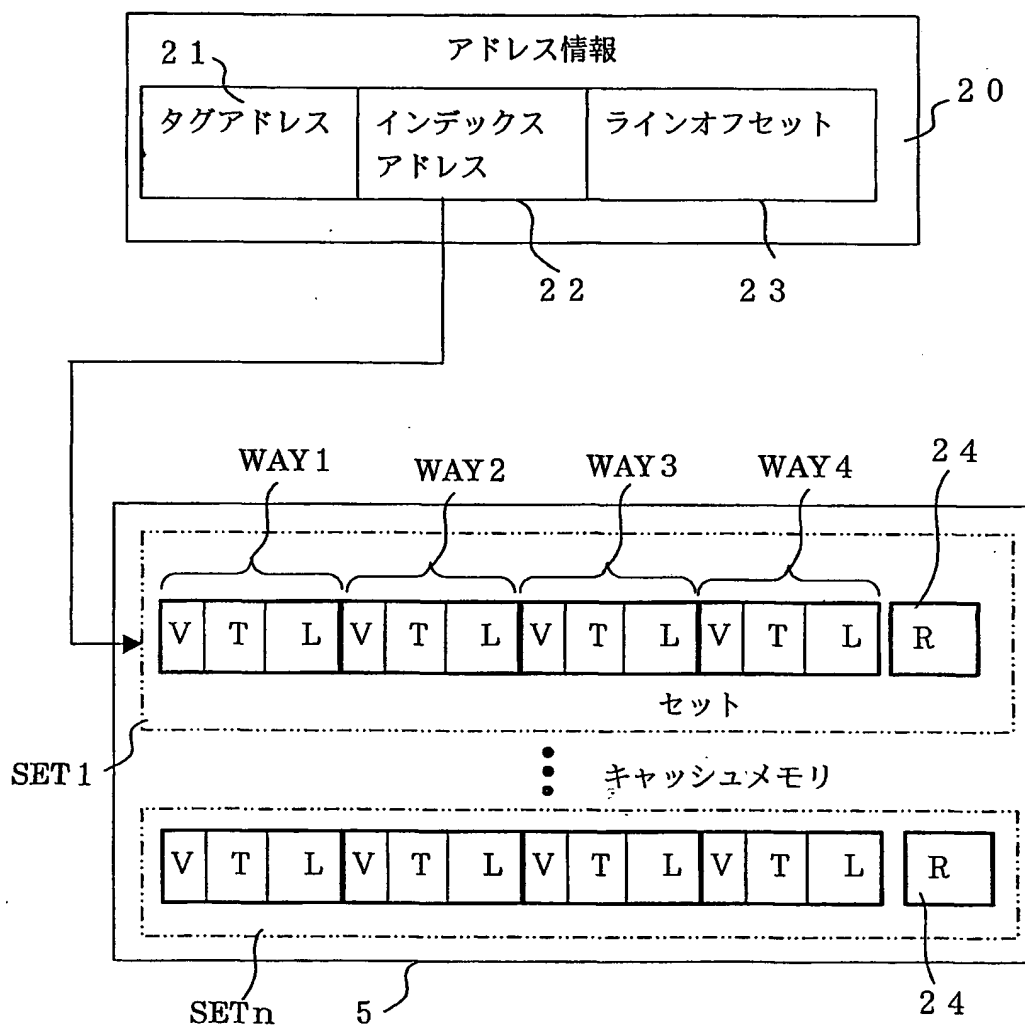
2/12

第2図



3/12

第3図



4/12

第4図

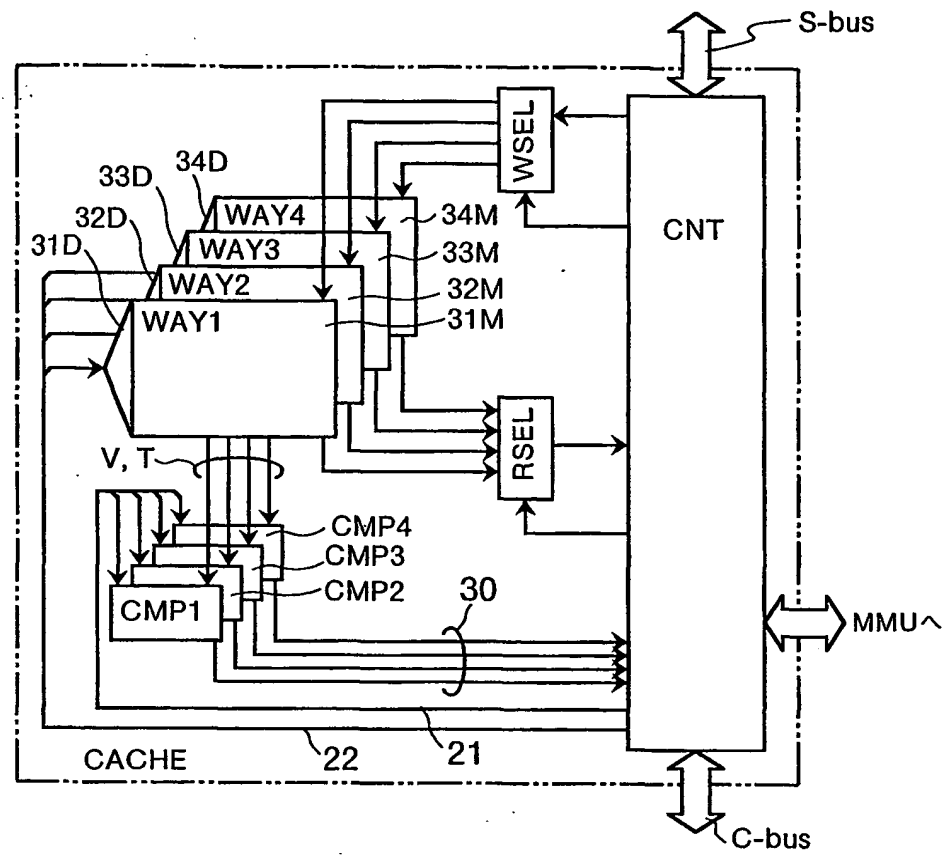
LRU 情報フィールド内 ビット位置	ビットの値が0のとき の意味	ビットの値が1のときの 意味
第1ビット	ウェイ1はウェイ2より アクセス時期が古い	ウェイ1はウェイ2より アクセス時期が新しい
第2ビット	ウェイ1はウェイ3より アクセス時期が古い	ウェイ1はウェイ3より アクセス時期が新しい
第3ビット	ウェイ1はウェイ4より アクセス時期が古い	ウェイ1はウェイ4より アクセス時期が新しい
第4ビット	ウェイ2はウェイ3より アクセス時期が古い	ウェイ2はウェイ3より アクセス時期が新しい
第5ビット	ウェイ2はウェイ4より アクセス時期が古い	ウェイ2はウェイ4より アクセス時期が新しい
第6ビット	ウェイ3はウェイ4より アクセス時期が古い	ウェイ3はウェイ4より アクセス時期が新しい

第12図

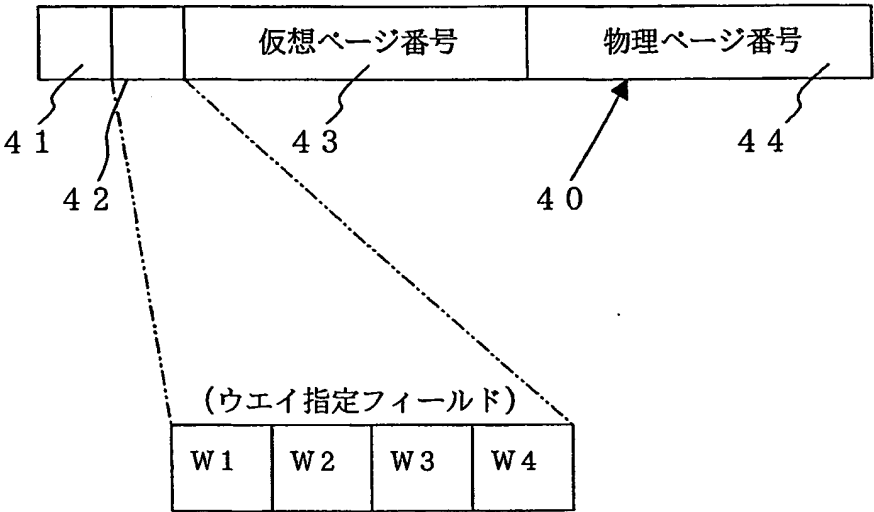
セクション情報名	意味
スタートアドレス	仮想アドレス空間でのこのセクションの開始 アドレス
セクションファイルオフ セット	実行可能ファイル内でのこのセクションの開 始オフセット
セクションサイズ	このセクションのサイズ
基本ページサイズ	このセクションの基本ページサイズ
ウェイ指定情報	このセクションを構成するページのウェイ指 定情報

5 / 1 2

第5図

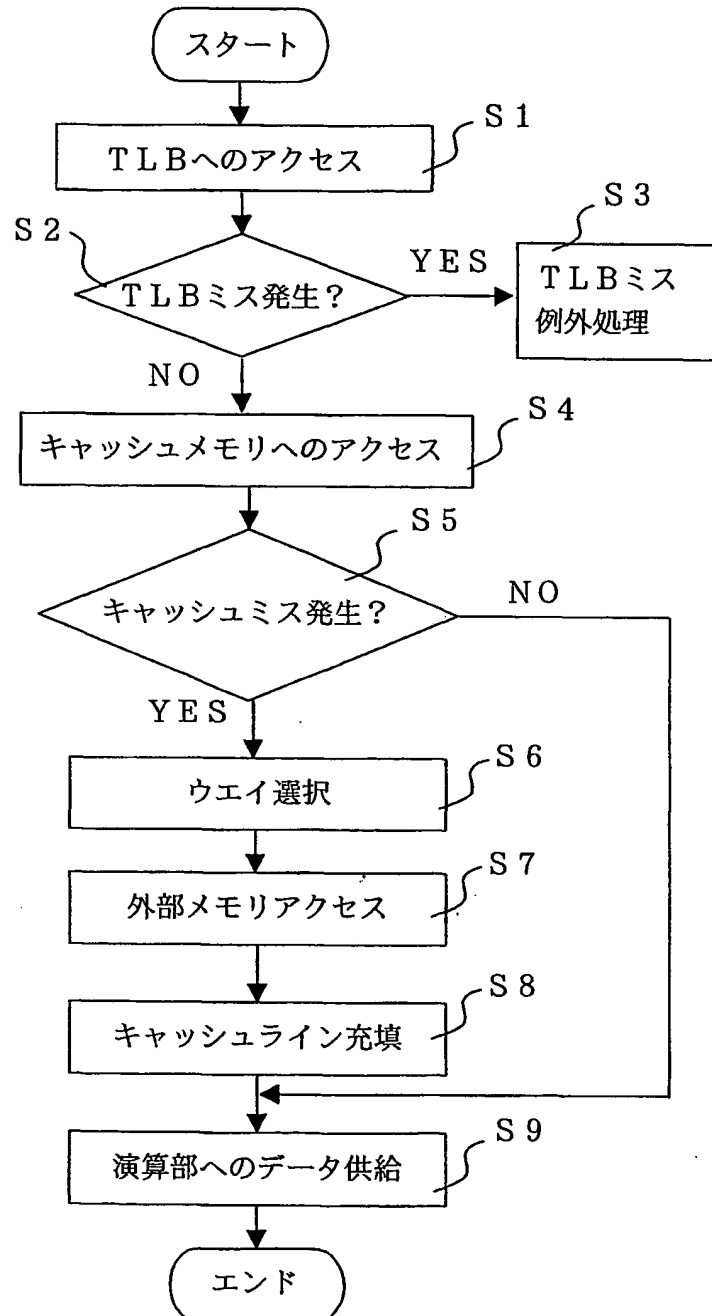


第 6 図



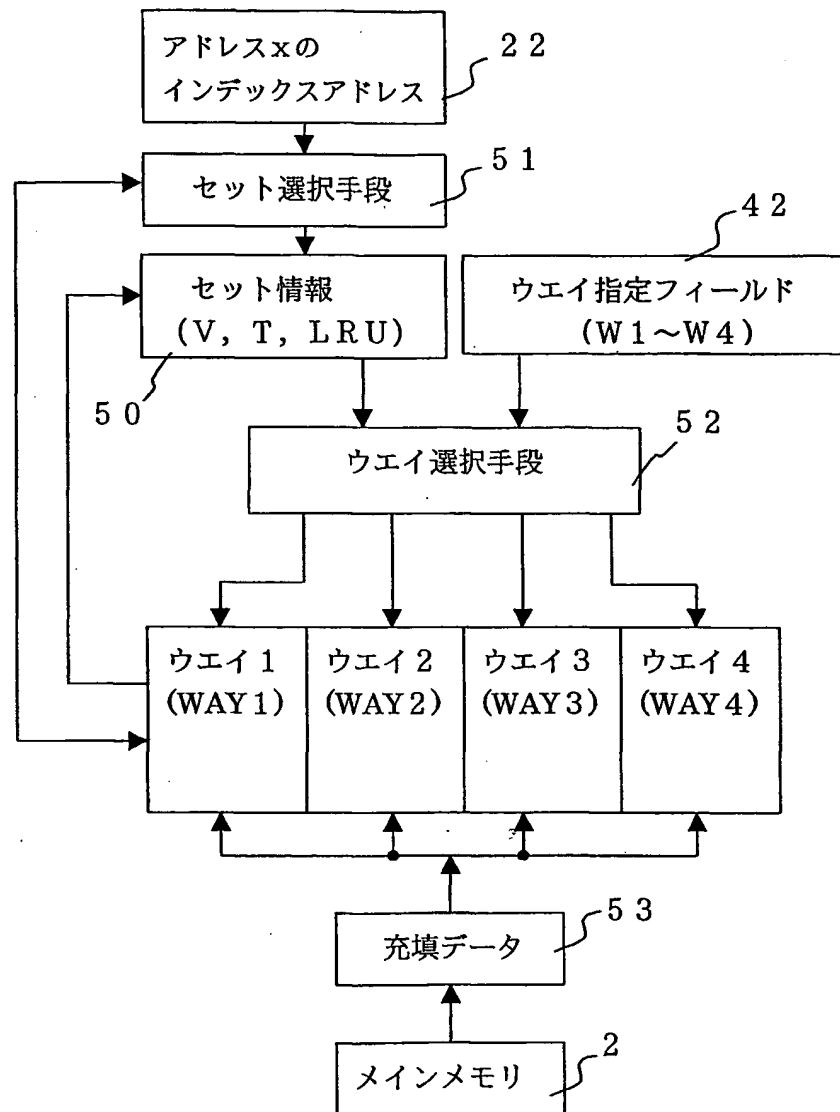
7/12

第7図



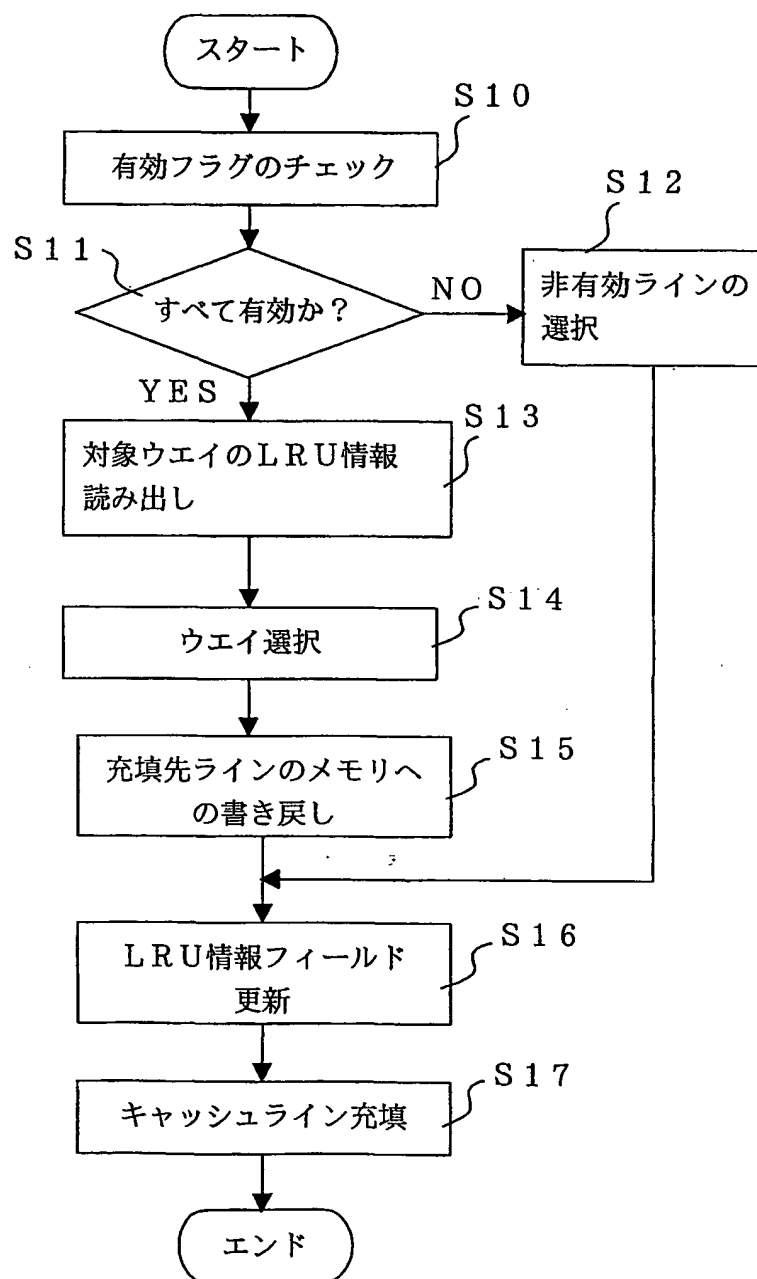
8/12

第8図



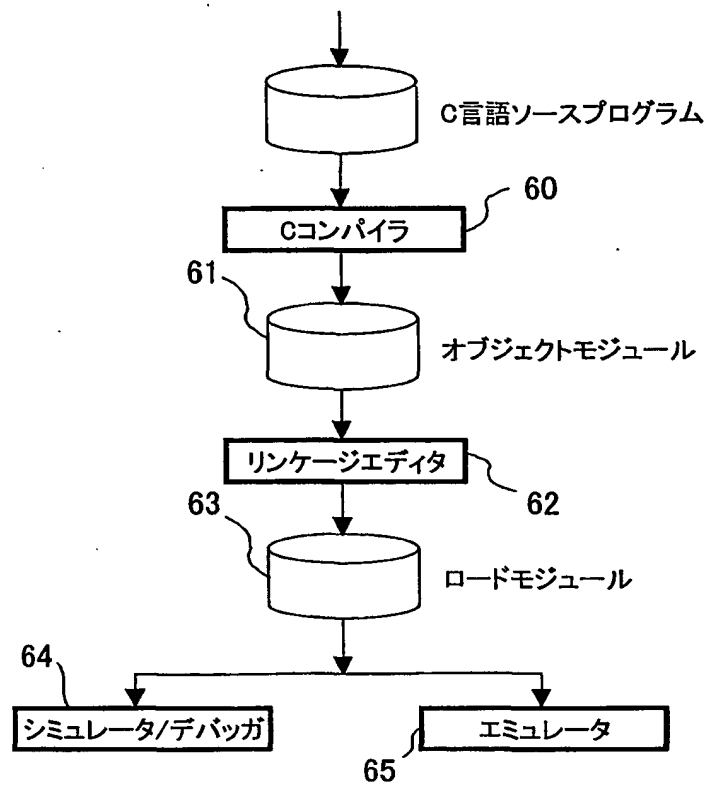
9/12

第9図



10/12

第10図



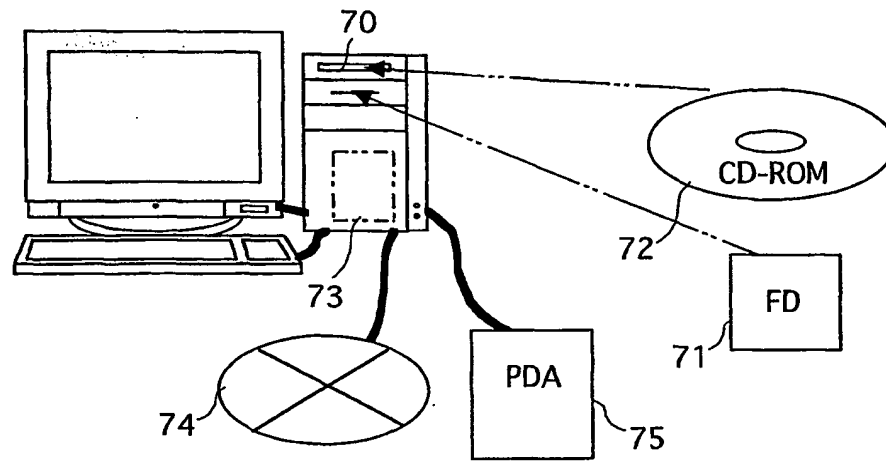
11 / 12

第11図

(a) #pragma section("シンボル名", "セクション名", "ページサイズ", "ウエイ指定")
(b) コード for(i=0; i<N; i++) for(j=0; j<N; j++) for(k=0; k<N; k++) C[i][j] += A[i][k] * B[k][j];
(c) コード (ファイル名: "code1.c") #include "libraryA.c" extern libfunc0; myfunc(...) { ... libfunc(...); }
(d) データに対する指示文 #pragma section("A", "dataA", "4K", "0100") #pragma section("B", "dataB", "4K", "0010") #pragma section("C", "dataC", "4K", "0001")
(e) 命令に対する指示文 ファイル"code1.c"内指示文: #pragma section("myfunc", "mycode", "4K", "1100") ファイル"libraryA.c"内指示文: #pragma section("libfunc,...", "libcodeA", "4K", "0011")

12/12

第13図



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP00/04906

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl.⁷ G06F12/12, G06F12/08, G06F12/10, G06F9/06, G06F9/45, G10K5/02, H04N5/907

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int.Cl.⁷ G06F12/08-12/12, G06F9/06, G06F9/45, G10K5/02, H04N5/907, G06F15/347

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched
Jitsuyo Shinan Koho 1922-1996 Toroku Jitsuyo Shinan Koho 1994-2000
Kokai Jitsuyo Shinan Koho 1971-2000 Jitsuyo Shinan Toroku Koho 1996-2000

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 856797 A1 (SGS-Thomson Microelectronics Ltd.),	1-14
Y	05 August, 1998 (05.08.98) & JP 10-232834 A & EP 856798 A1 & JP 10-232839 A & WO 98/34172 A1 & JP 11-509356 A	15, 16
Y	JP 4-49446 A (NEC Corporation),	1-14
Y	18 February, 1992 (18.02.92) (Family: none)	15, 16
Y	JP 8-212135 A (Hitachi, Ltd.), 20 August, 1996 (20.08.96), page 5, left column, line 30 to page 5, right column, line 2; Figs. 6, 7 (Family: none)	15, 16
X	JP 62-194564 A (NEC Corporation), 27 August, 1987 (27.08.87) (Family: none)	17-20
A	NEC Gihou, Vol.48, No.8, 11 September, 1995, Hiroki MACHIDA et al., "V850 Family muke Kouseinou C Compiler CA850", pages 42 to 47, especially, page 44	15, 16

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier document but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search
17 October, 2000 (17.10.00)

Date of mailing of the international search report
31 October, 2000 (31.10.00)

Name and mailing address of the ISA/
Japanese Patent Office

Authorized officer

Facsimile No.

Telephone No.

A. 発明の属する分野の分類 (国際特許分類 (IPC)) Int. Cl ⁷ G06F12/12, G06F12/08, G06F12/10, G06F9/06, G06F9/45, G10K5/02, H04N5/907		
B. 調査を行った分野 調査を行った最小限資料 (国際特許分類 (IPC)) Int. Cl ⁷ G06F12/08-12/12, G06F9/06, G06F9/45, G10K5/02, H04N5/907, G06F15/347		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2000年 日本国登録実用新案公報 1994-2000年 日本国実用新案登録公報 1996-2000年		
国際調査で使用了電子データベース (データベースの名称、調査に使用した用語)		
C. 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y Y	EP, 856797, A1 (SGS-THOMSON MICROELECTRONICS LTD.), 5. 8月. 1998 (05. 08. 98), & JP, 10-232834, A & EP, 856798, A1 & JP, 10-232839, A & WO, 98/34172, A1 & JP, 11-509356, A	1-14 15, 16
Y Y	JP, 4-49446, A (日本電気株式会社), 18. 2月. 1992 (18. 02. 92) (ファミリなし)	1-14 15, 16
<input checked="" type="checkbox"/> C欄の続きにも文献が列挙されている。 <input type="checkbox"/> パテントファミリーに関する別紙を参照。		
* 引用文献のカテゴリー 「A」 特に関連のある文献ではなく、一般的技術水準を示すもの 「E」 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの 「L」 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献 (理由を付す) 「O」 口頭による開示、使用、展示等に言及する文献 「P」 国際出願日前で、かつ優先権の主張の基礎となる出願日の後に公表された文献 「T」 国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの 「X」 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの 「Y」 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの 「&」 同一パテントファミリー文献		
国際調査を完了した日 17. 10. 00	国際調査報告の発送日 31.10.00	
国際調査機関の名称及びあて先 日本国特許庁 (ISA/JP) 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官 (権限のある職員) 清水 泰 電話番号 03-3581-1101 内線 3585	

C (続き) 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求の範囲の番号
Y	JP, 8-212135, A (株式会社日立製作所), 20. 8月. 1996 (20. 08. 96), 第5頁左欄第30行目-第5頁右欄第2行目, 第6図, 第7図, (ファミリーなし)	15, 16
X	JP, 62-194564, A (日本電気株式会社), 27. 8月. 1987 (27. 08. 87) (ファミリーなし)	17-20
A	NEC技報, 第48巻, 第8号, 11. 9月. 1995, 町田弘樹, 中村さおり, 森本真一, 西海弘恭, 「V850ファミリー向け高性能CコンパイラCA850」, p. 42-47 特にp. 44参照	15, 16